TSM-Bench: Benchmarking Time Series Database Systems for Monitoring Applications

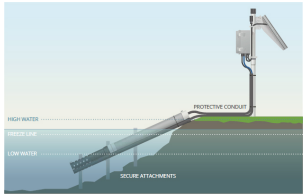
Abdelouahab Khelifati, Mourad Khayati, Anton Dignös,
Djellel Difallah, Philippe Cudré-Mauroux

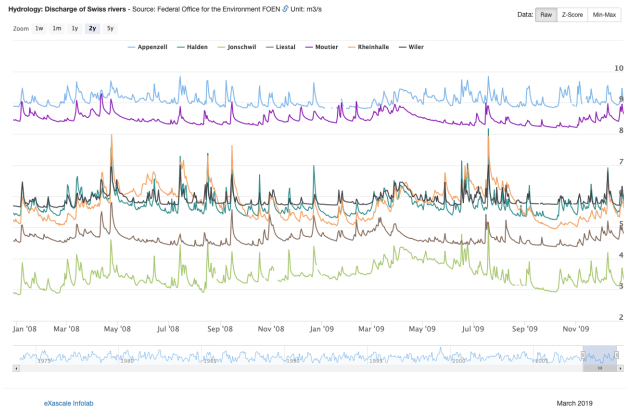
VLDB'23, Vancouver – Canada

August 30, 2023

Monitoring Application

- BAFU[1] use sensors to monitor the water quality in Swiss rivers.
- The collected time series are multivariate with different features.





[1] Federal Office for the Environment in Switzerland

Monitoring Requirements

- Monitoring hydrometric time series involves various analytical tasks: data exploration, anomaly detection, forecasting, trend analysis, recovery of missing values, and similarity search.
- Traditional RDBMs are ill-equipped to handle analytical tasks.
- Time Series Database Systems (TSDBs) are specialized systems that store, manage, and query large time-series data.
- Picking the best TSDB remains a challenge.

SOTA & Contributions

- Existing TSDB benchmarks implement:
  - Static queries on a subset of relevant systems.
  - Ingestion and querying in isolation.
  - Simplistic data generation, if any.
- TSM-Bench benchmarks seven popular time series systems by providing:
  - Dynamic query evaluation using offline and online workloads.
  - Realistic time series generation technique.
  - Recommendations for understanding and navigating the architecture of systems.

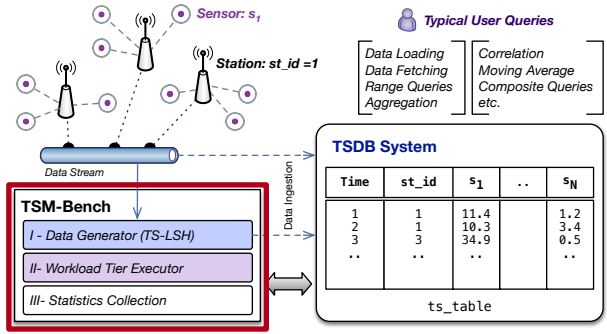
Current Section

TSM-Bench

Evaluation

Recommendation

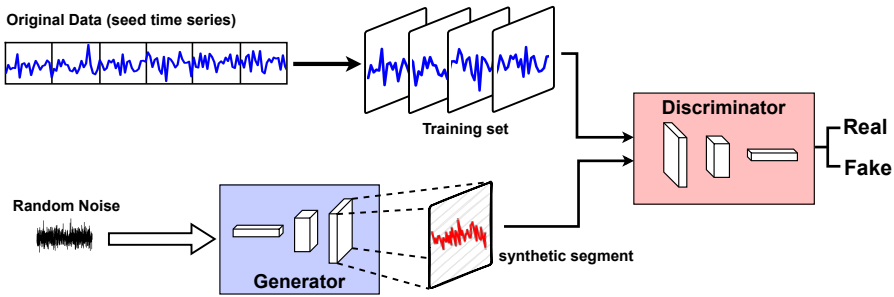
Architecture



- TS-LSH uses sample data to generate large realistic data streams.
- The executor launches configurable workload tiers.
- The statistics collection module records the performance of the TSDB.

TS-LSH: Time Series Generation/1

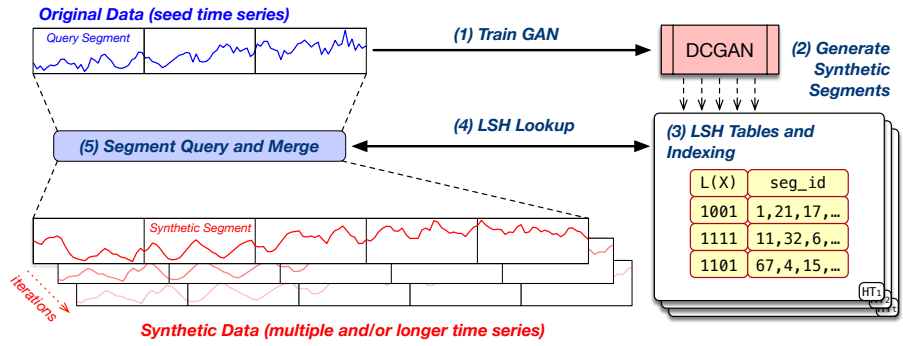
Generative Adversarial Network (GAN)



- GAN takes an input time series partitioned into segments.
- Concatenating segments using GAN is exponential.

TS-LSH: Time Series Generation/2

Locality-Sensitive Hashing (LSH)



- TS-LSH augments both the length and the number of time series.
- TS-LSH is sub-linear with the input and linear with the output.

Current Section

TSM-Bench

Evaluation

Recommendation

Workloads Design

We designed the workloads around three performance dimensions:

- Size of input/output data, data access, and the number of operations.
- Interplay between querying and ingestion.
- Impact of time series features on compression performance (data encoding).

Experimental Setup

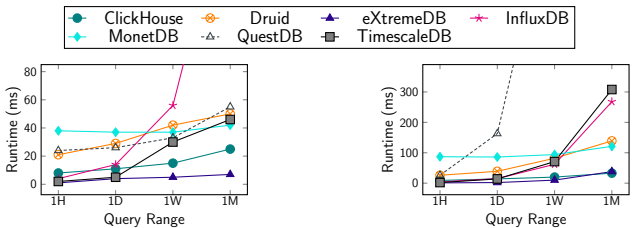
- **Systems**: ClickHouse, Druid, eXtremeDB, InfluxDB, MonetDB, QuestDB, and TimescaleDB.
- **Datasets**: Two datasets (#sensors, #stations, range)
  - D-LONG: $10 \times 100 \times 60$ (518M datpoints)
  - D-MULTI: $2k \times 100 \times 10$ (17.2B datapoints)
- **Offline workload**: 7 queries (Fetching, Fetching with Filter, Aggregation, Downsampling, Upsampling, Cross-sensor Average, and Correlation).
- **Online workload**: Queries execution under concurrent ingestion.
- **Compression workload**: Storage size under various time series features.

Offline Workload

Aggregation & Downsampling

```
SELECT st_id, AVG(s_i)...AVG(s_j)
FROM ts_table
WHERE st_id in <st_list>
AND time < ?timestamp
AND time >= ?timestamp - ?range
GROUP BY st_id;
```

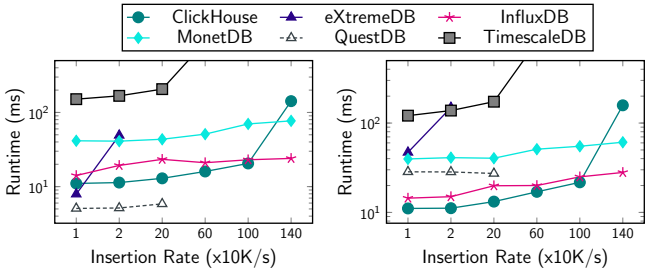
Agg



(a) Aggregation (b) Downsampling

- eXtremeDB and TimescaleDB are the fastest in case of simple aggregation.
- eXtremeDB and ClickHouse are well-suited to downsample data.

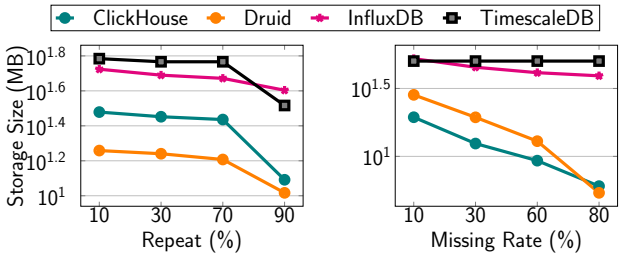
Online Workload



(a) Data Filtering (b) Data Aggregation

- Queries do not block writes for all systems.
- QuestDB and ClickHouse are the best for low insertion rates.
- InfluxDB and MonetDB provide the best runtimes for very high insertion rates.

Compression Workload



(a) Repeats (b) Data Sparsity

- All systems benefit from the existence of repeats.
- Only ClickHouse and Druid can take advantage of the existence of missing values.

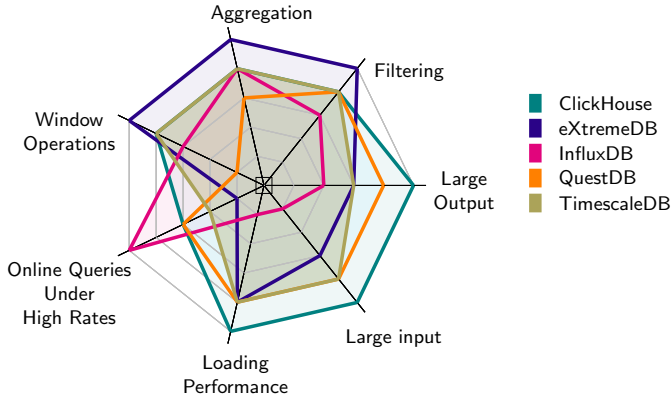
Current Section

TSM-Bench

Evaluation

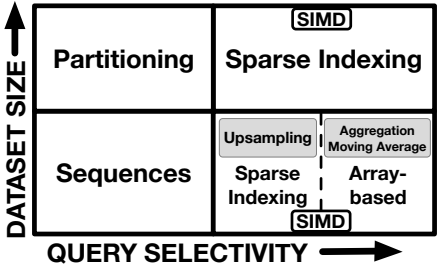
Recommendation

Performance Summary

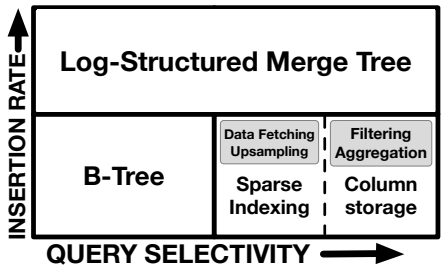


- Seven discriminative dimensions for comparing the performance of TSDB.
- The performance of each system for different query types is ranked on a 0-5 scale.
- No silver bullet system.

Architecture Impact



(a) Offline Workloads



(b) Online Workloads

- No single architecture dominates all the workload tiers
- Design factors:
  - Offline workloads: query selectivity and the size of the data
  - Online workloads: Insertion rate and the query selectivity

Conclusion

- TSM-Bench is a comprehensive benchmark for TSDBs.
- We provide a fine/coarse-grained recommendation for decision-makers at different levels.
- The code is open-source.
- Future work includes mixed-queries workloads and multitenancy scenarios.

TSM-Bench: Benchmarking Time Series Database Systems for Monitoring Applications

Abdelouahab Khelifati, Mourad Khayati, Anton Dignös,
Djellel Difallah, Philippe Cudré-Mauroux

Thank you!

abdel@exascale.info

Questions?

https://github.com/eXascaleInfolab/TSM-Bench

UNI FR

unibz

NYU ABU DHABI